Due to the large number of chemicals for which toxicological and ecotoxicological information is lacking, priority setting for data acquisition is a major concern in chemicals regulation. In the current European system, two administrative priority-setting criteria are used, namely novelty (i.e., time of market introduction) and production volume. In the proposed Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) system, the novelty criterion is no longer used, and production volume will be the main priority-setting criterion for testing requirements, supplemented in some cases with hazard indications obtained from QSAR modelling. This system for priority setting has severe weaknesses. In this paper we propose that a multicriteria system should be developed that includes at least three additional criteria: chemical properties, results from initial testing in a tiered system, and voluntary testing for which efficient incentives can be created. Toxicological and decision-theoretical research is needed to design testing systems with validated priority-setting mechanisms.

Key Words: chemicals control; risk assessment; risk management; REACH; tiered testing.

INTRODUCTION

There are few areas of applied science in which the road from science to policy is as difficult as in the risk management of chemicals. These difficulties depend in part on the lack of information about individual substances, in part on the unavoidable extrapolations in risk assessment, and in part also on the lack of agreed-upon principles for decision-making under uncertainty.

About 70,000 chemical substances are commercially offered on the European market. For most of these substances, the available information is not sufficient to make even an approximate estimate of the potential risks associated with use of the substance (Allanou et al., 1999). In recent years, major attempts have been made to remedy this lack of information. An ambitious initiative to increase data generation has been taken in the United States (the Chemical Right-to-Know Initiative’s high production volume program), in which chemical producing companies have voluntarily committed to provide basic information for over 2000 high-production-volume substances (http://www.epa.gov/chemrtk/). Furthermore, the European Commission has developed a proposal for a new regulatory system for chemicals control, the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) system, aiming at substantially increasing the knowledge base on which chemicals risk management is based (European Commission, 2003).

These initiatives will certainly improve the situation, but clearly this enormous knowledge gap cannot be filled in one piece, and there is an urgent need to limit the costs and the number of animals used for additional testing. Therefore, an improved strategy is needed to cope with the lack of data. Several components for improving test strategies have been proposed, including the use of exposure information, thresholds of toxicological concern, (Q)SARs, read-across methods, and in vitro testing (see e.g., Bradbury et al., 2004; Cronin et al., 2003; TAPIR, 2005; van der Jagt et al., 2004). Two main questions in these discussions are: Which tests and which substances should be given the highest priority? Can we make meaningful preliminary assessments based on simple initial tests or perhaps even on the chemical properties of a compound? There is an urgent need for priority setting criteria that can be used to set up test systems to serve the purposes of risk management as efficiently as possible.

CURRENT CRITERIA: NOVELTY AND PRODUCTION VOLUME

In the current European regulatory system, two administrative priority criteria for test requirements are used, namely novelty and production volume. The criterion of novelty takes the form of a division of substances into two groups, those introduced on the European market before and after September 18,
1981 (in the regulatory context, these are the “existing” and the “new” substances, respectively). The criterion of production volume is used to select, among the existing substances, those that should first be risk assessed, and among new substances, which tests to require. New substances are subjected to much more stringent testing requirements than “existing” substances (Council Directive 92/32/EEC; Council Regulation 793/93).

In REACH, the novelty criterion is abolished, and the same data requirements will be applied to new and existing chemicals (European Commission, 2003). Good arguments can be given for this step. We have no reason to believe that substances that have been synthesized or introduced on the market recently are in general more toxic than other substances. Furthermore, stricter requirements on new substances may possibly thwart innovation, so that new less toxic substitutes will not be developed.

Instead of the rejected novelty criterion, the proposed REACH system makes increased use of the production volume criterion. For existing chemicals in the current system, production volume is primarily used for temporal priority setting (“high-volume chemicals first”). In REACH it will also be used for priority setting with no time limit (“lower demands on low-volume chemicals”). This large-scale introduction of production volume as a criterion for stringency of regulation does not seem to be the outcome of careful analysis. The rationale for this criterion is—rather obviously—that large production volumes increase the potential for exposure and, therefore, also the potential risks associated with use of the chemical. This is a sensible argument, but the production volume criterion is nevertheless problematic in at least three ways. First, it is not known to what degree production volume actually predicts exposure. Second, there are indications that chemicals with low toxicity may be overrepresented among high-volume chemicals (Cunningham and Rosenkranz, 2001). This implies that the exclusive use of the production volume criterion could result in a priority-setting procedure that is biased toward less (acutely) toxic chemicals. Third, the lower exposure predicted for low-volume chemicals refers to aggregate (total) exposure, not to individual exposures. Exposure to a low-volume chemical may very well be (and often is) restricted to a smaller number of persons who may be exposed to large doses of the chemical, for instance, at their workplaces, even if the total production volume is low.

Data obtained by QSAR modelling (quantitative structure activity relationships) have furthermore been introduced for use as a priority setting tool in the REACH system as the proposal was voted on by the European Parliament on November 17, 2005. According to REACH Article 11, QSAR data providing “indications” that a chemical is persistent, bioaccumulating, and toxic (PBT), or very persistent and very bioaccumulating (vPvB), or causing CMR effects (carcinogenicity, mutagenicity or reproductive toxicity), can be used for triggering test requirements for low volume chemicals (1-10 metric tons per year). If such indications are available, Annex V testing can be required, that is testing for skin irritation/corrosion (in vitro), eye irritation (in vitro), skin sensitization (in vivo), mutagenicity in bacteria, and short-term toxicity to Daphnia. In our view, this kind of regulatory mechanism has the potential to improve priority setting. However, in REACH as currently proposed, the connection between the priority setting criteria, i.e. the QSAR indications, and the test requirements that follows need to be strengthened. In particular, the required tests should be more relevant to the indicative data at hand. The scientific rationale for requiring testing for skin and eye irritation, skin sensitization, mutagenicity in bacteria, and short-term toxicity to Daphnia for chemicals with indications of PBT, vPvB, or CMR properties is not clear. The appropriate QSAR models are furthermore yet to be specified.

TOWARD A MULTICRITERIA SYSTEM FOR PRIORITY SETTING

These weaknesses of the production volume criterion give us reason to look behind the current REACH proposal and investigate how priority setting can be optimized in the further development of the system. Due to the uncertainties involved, it is advisable to develop a system for priority setting that combines several mechanisms, rather than relying on a single criterion with uncertain properties. We propose that, in the further development of REACH, three additional priority-setting mechanisms should be introduced (see Fig. 1).

The Chemical Properties of the Molecule

The inherent characteristics of the chemical molecule determine its fate and behavior in the environment and

![Fig. 1. Three approaches to priority setting in chemicals control.](image-url)
consequently influence the potential for exposure. Chemical characteristics such as persistency (P) and bioaccumulative potential (B) can be used as predictors of potential exposure and as useful components in a first-tier characterization that can guide further testing and risk management. Therefore, it would be an important improvement to require a set of P and B data, enough to apply the P and B criteria for all chemicals regulated by REACH. This will make it possible to determine whether or not the substance is PB (persistent and bioaccumulative) or even vPvB (very persistent and very bioaccumulative) according to the criteria included in the current REACH proposal.

An assessment of P and B properties is essential already at first-tier and at low volumes, since these chemicals sooner or later end up in the environment (e.g., as waste material). This is potentially problematic, given the properties of these compounds, even though they may be handled safely with respect to direct human exposures and even though the ecological impact of small amounts may be limited.

Categorization of chemicals according to their P and B properties at first tier is currently being used by Environment Canada in accordance with the Canadian Environmental Protection Act. In this system, P and B properties are used as the basis for priority setting among 23,000 chemicals on the Canadian Domestic Substances List. The purpose of this categorization is to identify chemicals for refined risk assessment or for risk management decisions, not primarily to determine further testing (Anonymous, 2000; Environment Canada, 2003). Another example is the REACH system. In REACH, P and B criteria are used to identify chemicals to be submitted to authorization requirements. However, the P and B data are not an integral part of this test system either. The limitation with this approach is that it only prioritizes chemicals with known unwanted properties. The major challenge in chemicals control is to identify the chemicals that have unwanted properties among the huge number of chemicals that we know very little about. In our view, chemicals identified as vPvB should be subject to risk management decisions without additional testing. However, for chemicals that have P and B properties (but that are not vPvB) the use of P and B data should be expanded to serve as an important basis for determining further T (toxicity) testing. In this way we can obtain a solid basis for determining the toxicological properties of these chemicals and thus enable an assessment whether they are P, B and T.

There is a potential for further developing improved methods for characterizing chemical properties, for instance by using indicators of chemical reactivity (Green and Bergman, 2005). The available and proposed predictive systems are still underdeveloped with regard to biotransformation-dependent events. In addition, it would be interesting to explore further how chemical properties can be used for initial exposure estimation, for instance, used as predictors of the most likely route of exposure (inhalation, skin absorption or oral intake) and of the bioavailability of the compound (e.g., potential for absorption).

It should be noted that data on chemical characteristics are obtainable without animal testing.

Results from Lower-Tier Toxicity Testing

The principle of tiered testing implies that the outcome of lower-tier tests may have effects on further testing requirements. Examples of such test systems are the regulatory testing requirements for pharmaceuticals, which are divided into several consecutive steps. For industrial chemicals, this principle can be strengthened so that initial testing has a more direct impact on further test requirements.

The tests included in currently used tiered test systems are all carefully constructed according to scientific principles. However, for a tiered test system to be validated, it is not sufficient that each individual test is scientifically sound. In addition, the test system as a whole, i.e., the combination of the tests and the rules for how tests follow each other, has to be scientifically validated. This is a weak point in current test systems. Whereas scientific principles have been used in the construction of individual tests, the systemic level is largely based on more intuitive judgments. At least two regulatory processes have recently contributed to putting focus on this deficiency. The first process aims to increase testing of previously untested chemicals (as reflected in the U.S. Chemical Right-to-Know Initiative and in the development of the REACH system). These initiatives create new demands on test systems, since they require the system to handle testing of a huge number of chemicals. The other process is the general intention to replace animal testing with nonanimal methods. This procedure involves validation of alternative methods per se, but also a validation of combinations of different methods in order to refine, reduce, and replace animal testing (see, e.g., Hartung et al., 2004). In order to fulfill the objectives to increase knowledge about previously untested chemicals, a scientifically justified and resource-efficient process for testing and risk assessment is needed. This includes the availability of reliable and relevant test methods (including alternatives to animal models), as well as scientifically validated testing strategies, that can be used at a reasonable cost. This is a significant challenge in its own right, and the need for further research activities in this area is correspondingly urgent.

Two major approaches can be used to improve this situation, namely a mechanistic and a decision-theoretical one. Mechanistic knowledge can provide us with well-founded presumptions on whether or not test outcomes should be regarded as independent of each other. If two tests give indications of different effects that are independent of each other, a result from one of the tests does not give a prognosis of the outcome of the other test. Therefore, in this case data from one of the tests give us no reason to refrain from
applying the other test. Such tests should preferably be arranged in parallel. In contrast, if two tests give indications of the same effect, results from one of them may give an indication of whether or not the other test is needed. Hence, positive results from mutagenicity tests would in most cases increase the need for tests of carcinogenicity. When tests are related in this way, they should, in general, be arranged consecutively in a test system, and it may be adequate to use one of them as a screening test to determine whether or not the other should be performed.

The decision-theoretical approach to the design of test systems should be based on information obtained from studies of the correlation between different test outcomes and, ultimately, between these outcomes and human or ecosystem risk. However, in most cases it is not possible to determine the relevance of a single test or a test system in relation to effects in humans or entire ecosystems. Instead a reference test, such as a well-established long-term in vivo test, often has to be used as the gold standard (Hartung et al., 2004). Validation can also be made against a well-established combination of tests. Statistical information from the tests that have been performed can be used to determine the predictive power not only of individual tests but also of test systems that combine several tests. In our view the best result can be achieved by combining the mechanistic and the decision-theoretical approaches.

Incentives for Voluntary Testing

According to the REACH proposal, test requirements for the about 20,000 chemicals produced in one to ten metric tonnes per year are severely limited. For none of these chemicals will the proposed data requirements suffice as a basis for hazard- or risk assessment. Even if the REACH test requirements were improved also for the low-volume chemicals, it will, for practical and economical reasons, not be feasible to test economically minor chemicals to the same extent as the most investigated chemicals. This implies that, even with the implementation of REACH, there will still be chemicals in commerce for which uncertainties remain regarding their inherent properties to cause harm. In our view, an efficient system for chemicals control should therefore include mechanisms that encourage manufacturers of chemicals to obtain knowledge also beyond what is legally required, and to make this information available to the users of their products.

The current regulatory system is so designed that the information obtained from voluntary testing (i.e., testing in addition to the mandatory tests) can increase regulatory restrictions on the use of a chemical (typically by adding warning labeling to the product), but cannot—with very few exceptions—lead to reductions in regulatory requirements (Hansson and Rudén, 2003). This will not be changed in the REACH system.

One way to increase the incentives for voluntary testing is to introduce an additional dimension into the classification and labeling system, namely the dimension of (eco)toxicological ignorance, coupled to a new classification category, “insufficiently investigated.” Chemicals classified in this category should be assigned a warning label, including a warning symbol, such as a question mark, to indicate potential danger. See Figure 2 for an example of what the new labeling may look like (Hansson and Rudén, 2003). If manufacturers can avoid this labeling by voluntary testing, this will create incentives to perform such testing.

Classification as insufficiently investigated should be based on a minimal list of toxicological information and applicable to all commercially available chemicals, regardless of production volume. A chemical should be classified as insufficiently investigated if the information required on that list is not available. The OECD Screening Information Data Set (SIDS) can be used as a starting point in discussions on what data should be required for not classifying a chemical as insufficiently investigated. However, it is essential that the criteria should be flexible and allow for different scientifically acceptable ways to obtain the required information. Such flexibility is necessary to avoid testing that is costly in terms of resources and animal welfare without providing information that is needed for the assessment and classification of the chemical.

The requirements should, furthermore, be considered for revision at regular intervals.

CONCLUSION

The regulatory system for chemicals control is inherently complex, taking a large number of aspects into account, and the actual outcome of implementing a new system for chemicals control in Europe has to be the result of a joint effort by different stakeholders such as the industry, the different member states, the NGOs, and scientists.
The above proposals for priority-setting criteria are preliminary and need to be further adjusted and amended. In order to develop and validate more efficient testing systems, we need improved basic mechanistic understanding of toxicity and exposure. We also need improved general knowledge about the prevalence of different inherent properties in the chemical universe as well as in defined classes of chemicals, and we need to establish datasets and test system that can be used for validation purposes (Hoffmann and Hartung, 2005). Research aiming at validating test methods has to use a different methodology than case-by-case mechanism-based toxicology. It requires statistical analyses of large datasets of compounds submitted to various tests. The analysis and interpretation of such data requires close cooperation between chemists, toxicologists, ecotoxicologists, statisticians, and decision theorists. The aim is to develop efficient, proactive, and scientifically validated test systems for the purposes of risk management. In such systems, priority-setting mechanisms should be integrated as first tier to determine further testing.

SUPPLEMENTARY DATA

Supplementary data are available online at http://toxsci.oxfordjournals.org/.

REFERENCES
